specificity, I find that one a tough one to handle.

What I would point out about moving along such a line is that if you start at this point and you move along a line that connects you straight to the top right corner, what you are doing is flipping a coin. That's the chance line from this point. anything that's a straight line up to the right-hand corner.

Now, the different ones fell in different places and we'll come back to that, but if you bear this in mind, a slide along this point is no change in positive predictive value, and again, when you look at my line box diagram, what that means is positive predictive value, to put it in graphic terms here is of all positives, how many of them are true positives.

As you can see from this diagram, this is just out of the air. This is not the data because I would have to draw this line much longer, but it would be something like this is about 40 percent of all of that. So that the true positives represent 40 percent of all positives. So the positive predictive value here would be 40 percent.

When we deal with mammography, for example, we talk about for biopsy recommendations, what is the positive predictive value. We find that

it's roughly 20 percent. It ranges from 15 to 40 percent, but it's roughly 20 percent for biopsies, meaning that of all biopsies, about 20 percent of them have cancer and 80 percent do not.

So positive predictive value, along with sensitivity I find the most intuitive and easy of all of these statistics to deal with.

So when we say that if the possibility that it moves along the line of constant positive predictive value, what we're saying is that the increase in true positives is the same percent of true positives as the increase in false positives is of the false positives, meaning that you maintain the same ratio of this to that or this to all of that, the same thing, if you move along this line.

And as you saw from Dr. Kondratovich's slides, some of these statistics did move along this line. One moved more or less along something that was just above the line of constant negative predictive value, which probably, although we don't know, it could have put us under the ROC curve, and to lose ROC area is a no-no.

So let me go back then to, again, the first of the third. So what we see is there was a gain in positive predictive value, meaning that for

the increase in the number of cancers called, the number of non-cancers that were included in that was even fewer percentage-wise, and that's something very important to consider when you're discussing this this afternoon.

Bearing in mind, of course, that the prevalence in the trial was one in three; in other words, out of the 240 chest X-rays, 80 of them had cancer. When we deal with the real clinical situation, I'm not sure what the figures are, but I think there's two orders of magnitude difference there.

When you deal with a population in which you are looking for cancer, we know that for mammography it's roughly one in 140 to one in 200 that have cancer, and so the positive predictive value is -- in other words, your right-hand portion of this is much, much larger. The whole thing is expanded in actual clinical practice.

Now, when we look at location specific, that is, looking at the person's sensitivity and their average, based on here getting the correct mark on the film, it went, the point estimates, from 66 percent to 68 percent, not very large, and we don't know what the statistical significance of this is. We don't have

the error bars on this. We may be able to calculate these, but we don't have them available now.

And regardless, the positive predictive value fell, meaning that it went down along below this line. It didn't go up here. It went to the right of that line, and that meant that there was a large number of false positives along with the increase in true positives.

Now to go to the second reading compared to the third, we'll go through this again. For the non-location specific, that is, the ROC analysis that we're able to do, the gain in area under the curve for all of the cancers, for the small ones and for the priors, there was a gain for all three of those.

And remember that on a previous one there was not for the 18 priors. As far as non-location specific sensitivity and positive predictive value were concerned, again, for all the cancers it went from 72 to 78 percent. That was statistically significant.

If you used the second of the third reading for all the reasons that Dr. Wagner explained, the increase for the small cancers was from 67 to 74 percent, again, statistically significant, and while we don't have error bars on it, the positive

predictive value probably didn't change very much one way or the other.

As far as location specific where you get credit for a real hit, but again, as I point out, if you get credit for putting a cancer in the wrong place, if you get a CT scan at least, then you can correct that most likely, although the company gave the radiologist the choice of checking for recommendation biopsy, not just CT, but biopsy directly without a CT. Certainly we might want to say this device when used should always, if positive, be followed by a CT, but that's not in the indication for use yet. I mean, that's something for the panel to consider.

Again, if we go to location specific, it went from 65 to 68 as opposed to 66 to 68, with the first reading on all cancers. Again, we don't know whether this is statistically significant. There was a drop in positive predictive value, again, meaning that when you lose location, you actually have an increase in false positives that is excessive compared to the increase in true positives.

And finally, the question of improved training might have had significant effect on these results. We don't know how much.

Thank you.

CHAIRMAN GARRA: Okay, everyone. So to give you a chance to digest all of that information, including the panel members, we're going to take a lunch break.

We'll start promptly at one o'clock, and there is no closed meeting today for the panel members. So we'll just have lunch.

So I'll see everybody at one.

(Whereupon, at 12:04 p.m., the meeting was recessed for lunch, to reconvene at 1:00 p.m., the same day.)

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

(1:04 p.m.)

CHAIRMAN GARRA: Okay. Welcome back, everyone. We're about ready to get underway.

What we're going to do is we're going to change the order of the meeting slightly because Dr. Ron Khazan was tied up in traffic in Washington, and believe me, I know what that's like. He's since arrived, but would like to leave before the snow hits. We all would, but he's got special dispensation here.

So anyway, we're going to let him speak now, and then we'll launch into the panel discussion with Dr. Toledano taking over.

So, Dr. Khazan, are you ready?

DR. KHAZAN: Yes, thank you.

I just wanted to emphasize -- I know you saw some of my slides before, but the three factors that make something like this important is the epidemiology, the prevalence of lung cancer, and the many deaths that occur because of it.

The advances in radiology that have occurred in the last five years, really more specifically with advances in CT, with spiral and multi-slice CT, we can now look at solitary pulmonary nodules much more specifically. We can evaluate them

better. We can follow them up more effectively, and we're in a much better position to deal with small pulmonary nodules.

Also, the computer and digital imaging advance make our ability to evaluate four algorithms to detect these on computers, and digital imaging would also be used as an input for this kind of computer program.

I think these three put together make this an apropos time for aiding the radiologist.

One thing I wanted to emphasize again was that the chest X-ray really is a very difficult, very busy film. I heard that mentioned before, and in general, we're not looking for cancer. We're looking for hundreds of other things. We have to look at the lungs, the air, the bones, and a solitary pulmonary nodule is more of an incidental finding when we see it.

I use this system, and I think the best way to describe it is like a medical student standing behind you. He doesn't know a lot about medicine, and he has not seen too many X-rays, but he's got a very keen eye. In other words, this points out lots of false positives. This system will take a normal film and show you three or four regions to look at.

But any radiologist with experience can dismiss with ease all of the false positives. Where this system is nice, in my experience is that it can look behind shadows that we may not see. It can look in the overlap of the right -- right under the right hilum and the right pulmonary venous confluence, and that's an area that most radiologists dismiss. Any density there is very likely to be confluence of shadows, but this system will look at that area and evaluate for a rounded density.

So it almost can look deeper in. Maybe it has better gray scale resolution even than radiologists looking at a chest exterior for hundreds of other things.

I think that utilizing this system over time would be much more effective. I think the radiologist that uses it day in and day out may learn its abilities and learn to dismiss its silliness and be able to use it as really a companion, pointing out hard to see areas.

I think the positives really outweigh the negatives.

Now, I didn't see the whole presentation, just the last two statistically oriented, and I have a couple of ideas about that.

There was a lot of discussion about specificity and positive predictive value. Let me give you an anecdote. Ten years ago, five years ago, before MR, anybody that came to the ER with a motor vehicle accident had a deceleration injury, got aorta squirted, arteriography. Why? Because we wanted to make sure that there was not a tear of the aorta, which would be a terminal event.

And we were able, we allowed ourselves one percent positive rate for all the morbidity of an arteriography of the aorta. In other words, we did that invasive procedure to 100 people hoping that we could catch one so that we could save his life.

Now, I think when looking at lung cancer, sensitivity is much more important than specificity. We are in a day where CT is used all the time. CT is as common. People are considering screening CTs.

So if you have a procedure that increases the sensitivity and all of those people go to a CT, you have picked up many more cases.

I am willing to CT hundreds of people to catch literally a few dozen more cancers. So if the problem is specificity, which is the negatives are noted by the system, if that goes down, what's the significance of that clinically? I don't see it in

the paper.

And there was another. It was emphasized about the location. The study did not mandate putting a location down, and I think a lot of people -- I know I did. If I saw a very suspicious area, I may have recommended a CT knowing that it would find the area.

If there were multiple areas, you picked one, or you didn't pick any. You said, "This guy needs a CT and the cancer will be found."

Also, only one location could be chosen. So maybe the location data, the problems with it were stubborn radiologists that went with their first hunch despite what the computer showed them, or someone didn't put a location down, or there were multiple questionable locations. I think all of those are possible.

There was another point made. I don't know how important this is, but I would say in the clinical practice we never biopsy before we CT, and you know, a note was made of if this is approved, maybe we should only CT after it and not biopsy.

That's moot. That's not an issue, I think.

And also, the locations that this shows that a radiologist does not see, I hope and I assume,

are likely to be subtle. So going to biopsy directly is not an issue.

Again, I used it. I think it could in time be as a friend looking over your shoulder pointing out lots of silly things, but once in a while picking up a gem, and if that increases CTs a little bit, even negative CTs, that's fine.

What we're trying to do is find early cancers and more of them and questions.

CHAIRMAN GARRA: Okay. Thank you very much.

At this point, now we will proceed with the open panel discussion. This discussion will be led by Dr. Alicia Toledano, who is the lead reviewer for this PMA.

So I'll yield control of the meeting to Dr. Toledano.

DR. TOLEDANO: Dangerous, dangerous, dangerous.

So my name is Alicia Toledano, and I would like to thank the FDA for the opportunity to be the lead reviewer on this PMA, and I'd like to congratulate the sponsor for putting together a very comprehensive application.

I have five and a half pages of questions

on your application that I've given to Bob Doyle, and they will be forwarded to you as well, I do believe, a six page or seven page summary and then five and a half pages of questions.

And the good news is that some of them have actually been answered this morning, the going to biopsy or not being the first, you know, one of the most important.

Many of my concerns are the same as those raised by the members of the FDA. I also have concerns about generalizability, about clinical relevance of the reading conditions, about false positives, about the results for the actionable priors.

And I know that my fellow panels members have concerns as well. So what I would like to do is open up the discussion for hopefully a very vigorous and participatory discussion by the panel, and we'll just keep going and let things fly,. and then after about an hour or so, we'll try to narrow in on some of the discussion points.

So who wants to ask the first question or raise the first issue?

DR. BERG: I will.

DR. TOLEDANO: Dr. Berg.

DR. BERG: Hi. Dr. Wendie Berg.

I guess the overwhelming question to me is really even with the contributions of this device, we're not looking at stellar sensitivities, and I think the question that I need answered the most is: is this really a clinically relevant device in practice or should we really be doing CT for screening for lung cancer?

DR. TOLEDANO: Do any other members of the panel have similar concerns or ideas about that? Dr. Mehta?

DR. MEHTA: Yeah, I would like to expand the question a little further. I think it's the same question, but in a broader sense.

The whole day has been filled with a lot of acronyms, and I would like to add my own. We've had a lot of PPVs and all of that. I would like to add a PSV, a positive societal value, because that's what I'm really baffling with here.

I think there is little doubt that this device very minimally improves the additional new lung cancer patients that can be picked up, in an enriched population where one-third of the patients we know already have cancer, also in a population where we're not told about the others, whether they were high risk

for cancer or not.

Now, let's take that scenario and transplant that to the statistic that was presented to us this morning. Approximately 60 million chest X-rays done in the United States every year. Let's put these 60 million chest X-rays through this device. We pick up on average five new nodules for X-ray. That's 30 million nodules.

Two minutes per radiologist for a nodule, that's 60 million minutes. How many new radiologists do we need to assess this? And how many new cancer patients will we pick up?

At the end of the day, for every lung cancer patient we pick up, what is the cost? And I would like to see some statistical analysis and a cost-benefit ratio perspective to address that. For every new lung cancer patient we pick up, what is the cost of this?

DR. TOLEDANO: Are there other members of the panel that -- go ahead.

CHAIRMAN GARRA: I'd like to also try to broaden that a little bit also by asking the FDA to please comment on they do require a cost-benefit analysis. We saw it in the PMA, and what are the components of it and how are they weighted in an FDA

decision?

I think the committee could use that information because the cost-benefit analysis that I saw left out some key ingredients, and we want to know is it worth pursuing or is it what you instructed them to do or what?

DR. TOLEDANO: Is there anybody from FDA who would like to answer that?

MR. DOYLE: Mr. Segerson?

MR. SEGERSON: Let Dr. Sacks address that.

DR. SACKS: Yeah, Bill Sacks.

We don't evaluate cost-benefit when we're looking at devices. We look at risk-benefit, but we don't evaluate cost-benefit. So, therefore, we didn't give you those figures.

DR. TOLEDANO: That's a very nice, concise answer.

Did we have further questions or clarifications or elaborations from the panel on this issue of clinical irrelevant improvement and cost per patient picked up and requirements of cost-benefit analysis before we maybe ask the sponsor for a two-minute answer?

CHAIRMAN GARRA: Let me just comment. In light of that answer, if cost to society and cost

monetarily or whatever is not an issue that the FDA is going to use in its deliberations, then we're sort of left with then we have to use it in the labeling.

In other words, if we were showing a very small improvement, I think maybe it would have to be reflected in the labeling or if it was a large one it might have to be reflected, or at least we could recommend that. What the FDA does, who knows?

DR. TOLEDANO: Dr. Smith?

DR. SMITH: If I may, I think it's a laudable goal to increase the sensitivity of detection of these lesions, and I think we're all in agreement about that. I think really what we're talking about is the clinical significance of this device, and at least in my way of thinking, when you've got a population -- I think it was 240 films and 80 of those films had cancer with them -- and you have only a small increase in sensitivity, I wonder about what the efficacy/effectiveness of this will be in the general population.

Along those lines, I agree with my fellow panel members that perhaps labeling, being very clear up front what the sensitivity of this device is might be appropriate.

In other words, I think a blanket

statement that this increases the conspicuity of lung lesions might be a little bit I don't want to say excessive, but something that folks could be reading and making one conclusion whereas the numbers suggest something else.

DR. TOLEDANO: Thank you.

Dr. Harms.

DR. HARMS: Yes. I just want to point out I don't believe we've been given a cost for the instrument. So how are we going to assess cost-benefit? We're not really charged with that task of cost-benefit.

The other issue is how much time does it take for a radiologist, which indirectly is a cost, and it would be helpful to get an idea of what kind of time commitment this is for radiologists.

I see the down side risks of this are the false positives, which would probably lead to more CT scanning, and we're not there dealing with a hazardous event. If you're talking about false positives leading to biopsy, directly to biopsy, then there would be significant risk in the false positives, but I don't see that really happening. I agree with the testimony on that.

So this has very little down side risk.

Potential up side of detecting more cancers. So it seems like a pretty good tradeoff.

DR. TOLEDANO: Okay. Before I allow the sponsor a minute to comment on the tradeoff, I also wanted to know how the conclusions of your cost-benefit analysis of one additional CT per cancer detected -- I think that was in the cost-benefit analysis. So it was stated that there would be one additional CT exam performed per cancer detected.

I wanted to know how did that conclusion depend upon the prevalence in the sample, those 80 out of 240, and how did it depend upon the particular operating point on the ROC curve or particular definitions of true positive and false positives and things like that.

So you all have what, 120 seconds to talk about clinical elements and improvement in doing CTs and these issues.

Dr. Freedman. Oh, hold on a second. Before you go, Dr. Mehta.

DR. MEHTA: I'm not sure that we even need to put dollar amounts on things.

DR. TOLEDANO: Right.

DR. MEHTA: I want to be clear when we talk about cost-benefit issues. It's not just dollars

that we are talking about.

For example, if one of us puts on a different hat in this room for a second, the hat of a hospital administrator, let me ask the question from that perspective.

I have the typical, absolutely average hospital in the United States, and I'm going to take 10,000 chest X-rays from my hospital and put them through this machine. Tell me how many more cancer cases I'll pick up.

DR. TOLEDANO: Okay. Dr. Freedman, go.

DR. FREEDMAN: Okay. Obviously I cannot answer all of those questions in 120 seconds. Let me give you some baseline on which to base a decision.

If you look at the Hopkins early lung cancer study from the 1970s, in their prevalence screen, which is equivalent to what we did here, they called back 25 patients for every cancer seen. So they called back a large number to find those cancers. That's the first thing.

The second thing is that in doing that, their average cancer size for the ones where they were sure it was cancer was 35 millimeters, and for the ones where they were suspicious that it might have cancer, it was 25 millimeters.

We're working with an average of 15 millimeters. So we're working with smaller cancers, and the population study based on 10,500-and something volunteers showed one cancer for every 25 call-backs.

Now, we don't know what the population effect of this will be. In an ROC study, you have to limit the size of the population to the study, but two things happen in that event. If you have a very high incidence of cancer and you tell the radiologist this is one specific task, they are reading with the maximum sensitivity that they could read under any circumstances you can imagine.

And if you tell them in addition, as we did, that the average radiologist picked up only two-thirds of the cancers in this kind of data set, they're looking as hard as they possibly can. They don't want to be embarrassed by the computer.

So even though we've seen only a small percentage increase in cancer, that's against the highest possible sensitivity that radiologists have themselves. We would expect that in a true clinical setting that the improvement in sensitivity would be even greater.

Now, is this cost beneficial? It would depend very much on the population that you're dealing

with. If you look at the study that Claudia Henschke reported, she had 23 cancers prevalent in 1,000 CTs. That's a very high instance, but she chose a high risk population.

If you choose a low risk population, then obviously your benefit is going to be a lot lower. This will increase the cancer detection rate. In routine clinical use, we do not know the percentage, but we think it will be a greater percentage than we've shown in the study.

DR. TOLEDANO: Okay.

CHAIRMAN GARRA: Brian Garra.

I'd like to -- what is that percentage? I saw several numbers scattered through the text, and I ended up being a little bit confused. I saw numbers as low as like four percent. I saw numbers of eight percent, 14 percent, 24 percent, and in various parts of the documentation.

DR. FREEDMAN: Fine.

CHAIRMAN GARRA: Which one is it going to be?

DR. FREEDMAN: Well, if you notice our claims, we do not claim a percent, and the reason is that percent depends on how you define it. If you define it based on whether or not the patient goes to

a CT, then you are very population dependent in any of these, but in that case, the overall improvement was as I remember nine percent or ten percent, and if you look at the population of nine to 15 millimeters, it was, I think, 15 percent or so. I don't have the numbers in my head.

Then you look at the alternative, which is you require exact location. Well, exact location is going to give you a lower sensitivity which is probably not clinically relevant because you'd get the CT if you identify the wrong location.

But to me that presents the lower bound of improvement. So what you get from the ROC area is the high level of what I would say is maximum improvement in this clinical trial. The one you get form location is the lowest benefit.

Now, in the location data, we did not design this study specifically to calculate that number. That was not the primary design of the study to use location to calculate sensitivity and specificity. We used location primarily to know what the effect was of the computer to give you these secondary analyses, which is what we used it for.

Many of these cases did have more than one signal or one lesion on the film. This is a standard

clinical population. It's not an experimental population where there's only one signal on a film to eliminate ambiguity. These cases, both the cancer and cancer free cases, are from a heavy smoking population. It's what you would see in a high risk population.

So I can't give you a precise percentage. I think it's closer to the ROC area percentage than the location percentage. It's somewhere in between.

DR. TOLEDANO: Okay. Let's let somebody else open up a new idea. I'm going to send you back to your seat.

DR. FREEDMAN: Good. Thank you.

DR. TOLEDANO: Thank you, Dr. Freedman.

DR. MEHTA: Can I ask a clarification question?

DR. TOLEDANO: Yes.

DR. MEHTA: I want to be certain that the applications are not asking for this as a screening test. That's the understanding I got, but as an adjunct to normal reading.

The reason I ask that is because if it were a screening test, you would limit it to a high risk population, and you would then compare it with things like screen CT, which is, in fact, a screening

test for a high risk population.

But if this is an adjunct, it's all comers across the board without screening for what the population is. Is that a correct interpretation? Is that what the application is asking?

DR. TOLEDANO: Let's have somebody from FDA give us the quick yes/no answer on that, Dr. Mehta.

MR. SEGERSON: Dr. Sacks, would you address that?

DR. SACKS: You tell us.

(Laughter.)

DR. SACKS: No, that's a very good question, and we would love for the panel to discuss that in terms of the labeling of the device and whether or not there should be a target population that is less than all chest X-rays.

DR. TOLEDANO: Okay. Thank you, Dr. Sacks and Mr. Mehta, for raising the question.

Dr. Segerson.

MR. SEGERSON: I thought it might be worthwhile looking at the indications for use again. Now, admittedly the one you saw earlier was already massaged a bit in the meeting we had with the company, and of course, we don't have a slide readily at hand,

but I think you have a copy; all the panel members have a copy.

But how do you read that?

DR. TOLEDANO: Let me just read it aloud, and while we're all looking for it, I will read it aloud -- Dr. Toledano -- and I would like to remind everybody to state your name when you begin speaking so that the transcriptionist can keep an accurate record.

The RapidScreen RS-2000 is a computer aided detection system intended to identify regions of interest on digitized frontal chest radiographs that may have features associated with solitary pulmonary nodules from nine to 30 millimeters in size, which could represent early stage lung cancer.

The device is intended for use as an aid only after the physician has performed an initial interpretation of the radiograph. Thus, the device assists the physician in identifying areas containing a potential lesion that previously may have been missed.

Now, first of all, did I read off today's slide?

MR. SEGERSON: Yes, the one that has been revised.

DR. TOLEDANO: Okay. That's great. Okay. So that's the current indication for use, and let's have the panel members discuss this.

Dr. Mehta, you raised the concern. Did you want to elaborate on the concern or would you --

DR. MEHTA: I mean, as I read it, this is not a screening tool. As I read it, this is applicable to the 60 million chest X-rays done in the United States annually. That's how I read the language as it's written.

DR. TOLEDANO: Dr. Garra.

CHAIRMAN GARRA: Brian Garra.

Dr. Mehta, is that something you agree with or --

DR. MEHTA: No, I do not agree with that. I don't think it should be used for the 60 million chest X-rays done. I think that's how the language reads right now.

The language, for example, does not say you should pre-select which patients a chest X-ray should be looked at, for example, you know, based on smoking history, exposure to risk factors for lung cancer, or anything of that sort.

CHAIRMAN GARRA: Oh, it's not specific. However, it does say, which could represent early

Stage 1 cancer. So at least that group would be targeted under this revised indication.

DR. MEHTA: But that's only on the basis of a chest X-ray finding, not a clinical history.

CHAIRMAN GARRA: Right.

DR. MEHTA: See, in any screening trial there's a clinical history. You pick a target group of patients, and then you screen them. I don't see that happening in this device indication for use.

No one says, "Go ask the patient do you smoke," for example. It's not our chest X-ray report. You know, the radiologists are sitting behind the room. They don't know whether this patient has ever smoked a cigarette in their life.

DR. TOLEDANO: Dr. Sacks.

DR. SACKS: Let me just say you're absolutely correct as it stands.

DR. TOLEDANO: Input from other panel members? Dr. Berg, Dr. Harms, Dr. Smith? Dr. Peters?

Nobody has any other?

I have some input. I guess when I read this, I naively assume that the physician who's performing the interpretation of the radiograph is looking at clinical history and is communicating with the patient's primary care physician and is making a

decision whether or not to use the computer aided detection device. So they're sort of ad hoc, separating out into a high risk population or not, and I guess the first thing I'm understanding is that that's not the way it works.

So if that's not the way it works, and I see two docs on the other side of the panel saying that's not the way it works, and if that's not the way it works, what do we need to say to make sure that it works appropriately? Do we have any suggestions for revised working? Do you have any ideas on this?

Let's take more panel ideas and then give the sponsor a few minutes.

CHAIRMAN GARRA: Brian Garra. I think that you're right because the person with all of the history is not going to be the one that calls the shots on whether the digitizer is being used or not. It's probably going to be in the radiology department, and we sometimes get some history, but we don't get the extent of history that you're thinking we might get.

I don't know that I would necessarily personally want to limit it to screening for primary lung cancer, although, and perhaps the manufacture could discuss this, they said there were many other

findings. It wasn't clear to me how many other real nodules there were that weren't cancer in their study, whether they selected so that they didn't have other real nodules or what.

What about metastatic cancer? What about all these other issues?

This also has a big effect on what you call a false positive versus a false negative, but I would think that it would probably pick up the metastatic nodule fairly well, just like it could pick up a primary nodule, and I wouldn't want to tie somebody's hands unduly, you know, but I do think having more information about what level of improvement they might expect would be useful than labeling.

DR. TOLEDANO: So it's Dr. Toledano again asking for more elaborations or concerns from the panel on this point before we allow the sponsor to state their perspective.

Dr. Berg.

DR. BERG: Yeah, Dr. Wendie Berg. I think one of the issues is that this really has only been validated in a high risk population. All of the chest X-rays that were used in this were all from patients who had at least a 20 pack-year smoking history.

So I think maybe either in the claims or in the labeling it has to be very explicitly stated that this has only been validated in the high risk population. That may be one way to encompass the concerns.

DR. TOLEDANO: That's an excellent point.

Dr. Toledano.

That's an excellent point, Dr. Berg.

Further elaborations from fellow panel members?

(No response.)

DR. TOLEDANO: Okay. Perspective from the sponsor? I recognize Dr. Freedman.

DR. FREEDMAN: Thank you.

Matthew Freedman.

It looks like I'll be the spokesman for a lot of these.

There are two problems. One is one can argue that screening should be done with CT first in the high risk population. The estimates that are published for that are an expense of 1.4 billion -- that's $1.4 billion in the United States in the first year, decreasing in subsequent years.

The problem is society cannot afford that as the screening method. Then you say who should be

screened with a system like this, and that depends on what you think the amount of money that you can spend on this screening process is.

CT is very expensive if you apply it to the whole population. The risk factors for lung cancer are well known, but they are graded, and I don't know and as a sponsor we don't know where a particular institution or physician will draw that line.

If you have a person who has smoked and who has COPD, they have eight times the risk of lung cancer as someone who smokes and does not have COPD. Therefore, maybe CT in the end will be limited to the very high risk people, and this will apply to lower risk people.

The other thing that is happening, and again, we don't know how to incorporate this, is the instanced of primary lung cancer in non-smokers is increasing. There is a very frightening study from Japan done at the Hitachi factory where 20 percent of the primary lung cancers were non-smokers.

So if we have to say how this should be used, we would say that this device should be used to screen for lung cancer in those below the very highest risk category, but not define a line because different practitioners may define that differently.

We think it should be used as a screen in people who have a known primary cancer, and you're looking for metastases in those situations where you would not use CT, and that is the person who is a year or two out from their primary cancer, not thought to have metastatic disease.

But the boundaries really depend on clinical practice, and we can define them somewhat, but not beyond that point.

DR. TOLEDANO: This is Dr. Toledano speaking.

Thank you, Dr. Freedman.

Dr. Mehta, you were about to make a comment.

DR. MEHTA: I want a clarification on the metastasis issue. First of all, we've seen zero data on metastasis. So I'm not sure that we can put anything in the label on metastasis in the absence of data.

Second of all, the lung is not the commonest site of metastasis for lung cancer. The commonest site of metastasis from lung cancer is brain, bone and other organs, not the lung, and I'm not sure in the absence of any data for screening for, you know, metastasis from lung cancer how we got onto

that discussion.

CHAIRMAN GARRA: Dr. Garra.

I think it was metastasis from other organs, not lung cancer metastasis.

DR. MEHTA: For which we have seen zero data today.

CHAIRMAN GARRA: We have zero data on both of those.

DR. MEHTA: Right. So I don't see how it can get in the label.

CHAIRMAN GARRA: Well, I don't think you want to put it in the label, but what you want to do is realize that people are going to try to use it for that because it is a nodule.

If you look at the labeling as it's stated here, and actually the problem with that sentence is it's very long. Lung cancer is at the end of a very long sentence, and most people run out of gas before they get to it.

But it really is saying it's intending to identify nodules that might be lung cancer, is what it's saying without all of the fluff in the middle, and I think that in a sense is a very appropriate label, in my opinion.

That's the intent. We're not forcing

people to use it just for that in this label, but that certainly that's its intent.

I guess I'd maybe try to rework that sentence a little bit though.

DR. TOLEDANO: Dr. Toledano.

Thank you, Dr. Garra.

As long as we have the device indication for use slide clearly visible to everybody, are there any other comments from members of the panel or concerns about these indications for use?

We've covered the early Stage 1 cancer in the first sentence. We've covered the second sentence about initial interpretation of the radiography. Any comments about identifying areas that previously may have been missed?

DR. BERG: Dr. Wendie Berg.

I'd like to -- I mean, it's a confusing issue because I think, as I understand their results, they did show significance comparing the CAD with the sequential read, but not with the independent read for that Claim 3, which is also now in the indication.

I'm not sure what to make of that, whether to then accept that as, in fact, a proven benefit or not because there's been a lot of discussion around that. I'd like to hear some more comments on that.

CHAIRMAN GARRA: Dr. Garra here.

I think the FDA folks have spent a lot of time thinking about that. Could we get a refresh from like you, Bob, Bob Wagner? No? About the issue of the sources of variability and whether in your experience, because you've seen a lot of studies like this and Alicia has as well, how strongly you feel that the lack of significant difference in the independent read -- how important that is.

DR. WAGNER: I think that's a professional issue. I don't think that I'm --

CHAIRMAN GARRA: But is it a show stopper? You know, sometimes you get a lack of significance and you're able to trace it to a specific problem, and it looked to me like you had, and thus, at least in both of your conclusions, it looked like you were placing less weight on the lack of significance in that because of that.

Maybe I was mistaken in my interpretation of that.

DR. SACKS: Bill Sacks.

No, you're correct in your interpretation. One of the things that I tried to say is that if you compare the second reading to the third and find a statistically significant increase, aside from the

randomness in the first reading compared to the third, it should be greater because there's less vigilance on the first reading than there is on the second reading which immediately precedes.

And that much less, if we were able, if we had the data, but it's an impossibility to compare it with clinical practice outside of trial, a point that I think Dr. Khazan made and perhaps Dr. Freedman. But I think that you have the sensitive probe with the second reading compared to the third, and it's valuable because it should be even less difference there than between the first reading and the third or clinical reading and the third.

And so if you can find it between the second reading and the third, by implication it's that much greater in practical terms.

DR. TOLEDANO: Do we have further comment?

Dr. Toledano speaking.

Do we have further comments from the panel members on this issue?

(No response.)

DR. TOLEDANO: No? Okay. I'll let loose on this one.

I think there are two issues that go into this question. The first is how you balance out the

different components of variability and standard errors of differences, and certainly based on the presentation by FDA this morning and some of the discussion in the PMA, I think we all appreciate by this point that there is less variability when you're looking at the difference between the sequential reads than there is when you're looking at the difference between the so-called independent read and the sequential read with the device.

So we understand that there's less variability, and many of us who are experienced with statistics know that if you're going to increase your sample size enough, you'll come up with significance.

I think the more important question is the clinical relevance and the clinical implications of the reading conditions because we're currently in a state in the field where we don't have CAD widely available, and that more readily or more easily could be associated with the independent read.

So when we're looking to say what happens if we move from the current state of lung cancer screening or chest x-ray interpretation to what would happen when we add in this device, I would think it's the comparison between the first and the third that makes more -- that's more clinically relevant.

I don't know if anybody else has perspectives on that that they would like to share. The sponsor has perspectives. Does anybody else on the panel have perspectives before we hear the sponsor perspectives?

CHAIRMAN GARRA: I would just say again I think you're right. It's probably more relevant, but there are some practical difficulties with analyzing that data which we've seen, and maybe when the sponsor gets up they can answer a question I've been wondering about, is the number of cases that they used was right on the borderline for significance. I wonder whether practical -- I mean there are a lot of lung cancer cases floating around in this country, and they mentioned even 10,000 of them in one of their slides. Yet we only see 80 cancers and 240 cases total.

Was that by design that they chose that number? Because as Alicia pointed out, if you go into larger numbers, you probably would have established significance on the independent reads, or may have.

So if you could address that as well.

DR. TOLEDANO: I'd like to thank Dr. Garra.

Dr. Smith has something to add.

DR. SMITH: Yeah, just one thing that I

guess in listening to this discussion I agree with you on the points that have been made.

It was mentioned earlier that the people would be that much more sensitive in the read just before the computer aided diagnosis. I wonder, too, just with the computer aided diagnosis, it is that, and when the computer identifies areas, regions of concern in a chest X-ray where you know a third of the patients have cancer or there's a high prevalence of the disease, is that isn't a little artificial, too.

And I guess just echoing my earlier comments.

DR. TOLEDANO: This is Dr. Toledano speaking.

Thank you, Dr. Smith.

I always have to remember to say my name at the beginning. I forget half of the time.

Dr. Freedman, did you have a response?

DR. FREEDMAN: Yes. This is Matthew Freedman.

There were two questions there. So the response to the first question is simply to keep in mind that the priors, these were missed prospectively by two radiologists. That means that those radiologists who knew that their primary task was to

detect cancer missed them.

Therefore, the real improvement in clinical practice should be even greater than what we showed because clearly they had shifted their sensitivity, given the format of the trial, to a much higher sensitivity for these cancers that had been previously missed by two people.

The second question is why did we use only 80 cancers. We used 240 cases because we calculated that that is what a person could do in half a day without fatigue, and that was the basic decision.

We also had a pilot study from which we calculated the number of cases that we needed. In the pilot study we had actually shown a much greater benefit than we showed in the larger clinical trial, and therefore, we felt that that sample would be sufficient.

As it turned out, it was not for both methods of comparison, but the sample size was based on that.

The second reason we did not use more cases is that in these very small cases, our sample cases are correlated, and so we had to make sure that only one film from patient was used, the current and the prior -- I'm sorry. The current and the prior

could not be used together.

And when you do that and you want to get this distribution in sizes, you end up with a shortage of cases.

DR. TOLEDANO: This is Dr. Toledano speaking.

Thank you, Dr. Freedman.

I do remember reading in the PMA that there were only 94 cases with lesions smaller than 30 millimeters from which to select these 80. I believe those were cancers that had been already through the quality control.

And so the difference between 80 and 94 is -- maybe could have been critical, but still there were only 94 available to begin with.

Did I get that right, Dr. Freedman? You can say no if I got it wrong.

DR. FREEDMAN: Again, it was 97.

DR. TOLEDANO: It was 97.

DR. FREEDMAN: And that was only if we did not use two films from the same patient.

DR. TOLEDANO: This is Dr. Toledano speaking.

Thank you, Dr. Freedman.

The ways that you can improve power in

these multi-reader, multi-case studies have to do with increasing number of cases, increasing the ratio of cancer cases to non-cancer cases, increasing the number of readers, and certainly reader burden. We all call this reader burden, the factor of how many films can a radiologist comfortably read during a trial because you are asking for favors from your colleagues when they come in and do these trials.

So I guess one question that I would have is you had 15 readers. Why not more?

Anybody want to answer?

DR. FREEDMAN: This is Matthew Freedman.

That was based on our pilot study and the estimate of the number of readers that we would need.

DR. TOLEDANO: This is Dr. Toledano speaking.

Thank you, Dr. Freedman.

I should just tape that and say, "This is Dr. Toledano speaking." Oh, that's wonderful. I've been wasting all this time. That's wonderful. My apologies for wasting all this time, and my thanks for bearing with me.

Dr. Mehta.

DR. MEHTA: This is Minesh Mehta here.

Since we're getting clarifications on the

patient population, again, this is more for clarification to make sure that the patient population that we are looking at in the study is comparable to the patient population today.

My understanding is that the vast majority of these lung cancer patients were selected from the old Hopkins database, which if I remember correctly excluded females.

The population in which this is disease is growing at the fastest rate in the United States today is females. The histopathologic distribution of this cancer in females is different than that in males. Its geographic distribution in the lung is different than that in males.

Even in males the histopathologic and geographic distribution of this disease has changed sine the Hopkins study. When you balance for all of those factors and you look at histopathologic and geographic distribution, can you tell us which specific histologies do you pick up more rapidly with this technology?

DR. TOLEDANO: Thank you, Dr. Mehta.

Actually before I allow sponsor to respond, I'd like other members of the panel to raise issues of generalizability and the characteristics of

the population to which this would apply.

Does anybody else have similar concerns? I know I had similar concerns.

CHAIRMAN GARRA: Brian Garra.

We did raise the issue about whether it might be applied to people with metastatic disease, and of course, then there's the people in the Midwest where they're going to have a lot of nodules anyway from granulomatous disease.

I don't know what that's going to do to your false positives, and I'm not sure whether you would call them false positives. This is not being marketed as a devise to distinguish benign from malignant nodules, but it's going to complicate things if you have instead of five because it picked up a rib shadow or something, if you have 50 or 60 circles on there.

Maybe the sponsors could discuss that briefly. They must have run into that on some of their cases.

DR. TOLEDANO: Are there other similar concerns from panel members?

Not yet.

Sponsor.

DR. FREEDMAN: This is Matthew Freedman.

We do have an overhead addressing one of the questions. So if that could be.

If you look at this on the screen, what we did, this is a performance test done with cases obtained by Deus Technologies. So these are not Georgetown cases, but these cases were separate from those used to train the system.

There were 98 men and 78 women. The sensitivity for the men was 68 percent. Sensitivity for the women -- this is machine detection -- was 66 percent.

You can see the average number of false positives per image is fairly similar. The Hopkins data set, indeed, is entirely male, and therefore, we did this study specifically to look at that question.

Now, the second question was about what would happen in the Midwest and in other places in terms of benign findings such as granulomas from histoplasmosis. If you look back to basically any study that's been done, but I'll use again the Hopkins study, but this applies to breast imaging as well; if you look at the Hopkins study, they had to call back in the prevalence screen 25 patients for every cancer found.

In the instant screen, because they had

the prior film available, they only had to call back nine.

Now, if you are working in a clinical situation, you are not working in the experimental situation. When you have a prior film, you will be able to recognize many of those granulomas as being granulomas because they've not changed. The system will alert you to them, and you will look.

The other part of Dr. Garra's question was what happens if there are 25 suspect regions there. The system uses a form of fuzzy logic to choose the most likely candidates for malignancy based on the criteria with which it's been developed.

So it will not give you 98. It will use various criteria. One of those criteria is a specific criteria to attempt to eliminate calcified granulomas. So, indeed, that has been taken into account by the sponsor.

Thank you.

DR. TOLEDANO: Thank you.

Dr. Mehta.

DR. MEHTA: I'm sorry to come back to my original question again because I'm not sure that I got the answer to my question from this overhead.

I understand that this is the sensitivity

of the machine for X-rays on either men or women. So it's imply measuring machine sensitivity. My question is this.

Women predominantly get adenocarcinoma. Adenocarcinoma are predominantly peripheral nodules. I could hypothesize, and this is simply a hypothesis on my part, that these might very readily be detected by the radiologists and that they do not need computer assistance in detecting this; that for all these women the computer adds nothing.

Prove me wrong.

DR. TOLEDANO: Sponsor.

DR. FREEDMAN: We have not done a clinical trial with women. We have only done a machine trial with women. We do have, though I don't have them here, the percentage of adenocarcinoma in the original setting of the cases that we used. There are a significant number of adenocarcinomas in that population. I do not have the precise numbers here, however.

I cannot prove you wrong that radiologists will routinely detect lung cancer in women without the computer. I would only point to the fact that there are several studies out there that show that in general, the detection of lung cancer in women on

chest X-rays is far inferior to the base detection rate in men.

We've shown that the machine detects them in women. We would expect that it should show a benefit in women in a clinical trial, but we have not done that clinical trial.

DR. TOLEDANO: Thank you.

More questions from the panel? No more questions from the panel?

Well, Dr. Garra has some more. Go, Dr. Garra, go.

CHAIRMAN GARRA: These aren't related to the -- we have spent a lot of time discussing the fundamental medical issues, but I wanted to just ask a few specific issues about the device itself. I'm looking for them here.

When I was reading the section on operation of the system, I noticed that you said in big, bold letters, I think, "Do not allow the UPS to shut this system down," and that struck me as a little odd since most modern systems talk to their UPS, uninterruptable power supply, and their UPS does a lot for controlled shutdown of a Windows 2000 system.

So you said the person must manually shut them down by turning off the power switch. Do you

have any of the engineers that could enlighten me on that?

The other question I had was this device has been developed in a pretty modern environment, and yet the user is required manually to type in information. It does not have a DICOM or HL-7 interface.

Those are the two questions I had.

DR. TOLEDANO: I'll continue actually with questions on technical specifications and general device issues before sponsor replies.

I just wonder basically what happens as technology advances. For instance, certain things happen in the Windows '95-'98 operating system, and the device now runs with the Windows 2000 operating system. What impact does that have?

What if you switched the laser printer? What if you get a better digitizer? What if you get a better monitor? What happens to the device as these things occur? I just want to know.

Dr. Smith and then Dr. Berg.

DR. SMITH: Along those lines, I wonder. At my institution we use a lot of computerized radiography and digital radiography is coming on line as well. How is that going to play with the system?

It looks like it's digitizing essentially not cut films, but printed images.

DR. BERG: That was my question as well.

DR. TOLEDANO: Okay. Further panel questions before sponsor replies?

CHAIRMAN GARRA: Just one other relating to the fact that you do digitize. Can you digitize a film version of an originally digital chest film and get a proper result or are we going to get aliasing and other problems?

Second, does the film screen combination have an effect on the performance of the system? Do you require certain specifications be met regarding the analog system you use?

DR. TOLEDANO: I think that's enough to keep you busy for a while.

DR. FREEDMAN: This is Dr. Freedman.

Let me just mention that there will be two technical people answering the question, I believe. One is Ed Martello, who is Chief of Engineering, and the other is Xin-Wei Xu, who Dr. Xu is chief scientist on this project.

DR. XU: I'm Xin-Wei Xu. I'm chief scientist in Deus Technologies.

Probably I can try to answer basically

ever you say what kind of case on digital director, digital or even digital, the image printed on laser film.

Basically we have already done an internal test for CR, which is from CR company, and we also did some tests which is just a drop of the film, which is a laser printed from another CR company, and basically and also for your question, we're using a laser digitizer. The test that we were doing in the lab, we didn't see too much difference.

Basically as I say, (unintelligible) take the digital imagine. As I come today what is our experiment to show basically it doesn't matter because I just brief you what the database we're training the system here.

We, as Dr. Yeh mentioned, we had collected more than 1,000 of chest images, which is actually from all over the world. Basically when we training this, we try to make our system can be adapt to any type of, kind of variety.

So imagines from different country basically had a variety range of exposure conditions, size, all of things. So I think our system is ready to go to any kind of digital data.

DR. TOLEDANO: Thank you.

Your next person?

DR. FREEDMAN: This is Dr. -- let me just comment. So what we did is we tested multiple forms of screen film combination, at least two different forms of digital data, direct digital and CR, and showed that the system, indeed, worked.

What we don't have is a clinical trial with that data.

DR. MARTELLO: This is Ed Martello, the engineering person to answer the part of the question that was engineering related.

In response to the shutdown of the system, we really wanted to prevent the user from going and going and going and possibly having a film being digitized and the system decide to shut down and leave a film in the digitizer.

We guarantee at least five minutes. We have tested it out to almost an hour of run time so that they should be able to finish what they're doing and shut down the system gracefully. That was the only reason.

DR. TOLEDANO: Thank you.

CHAIRMAN GARRA: Excuse me. Do you have a DICOM interface in the works or an HL-7 interface so that for use --

MR. MARTELLO: Yes.

CHAIRMAN GARRA: I'm thinking of this in terms of efficiency because it's going to take more time to read these, and usability of the system will be impacted by how much extra time it takes to do this, and if they have to type in the patient's name and everything again, people are going to have trouble with that. It may get under used.

DR. MARTELLO: We do have plans to add DICOM interfaces and things like that. There are real safety issues that we didn't want to address in our first product.

As soon as you get onto a network, you have all the issues of malicious use, viruses, and so on and so forth. Our engineering task was to get a device that was useful and eliminate most of that networked problems.

In addition, a lot of the industry is still film. So that was just a sequencing decision.

DR. TOLEDANO: Thank you very much.

Do other panel members have questions? I have another question I could ask, but I'll let anybody else ask a question first.

(No response.)

DR. TOLEDANO: No other questions? Okay.

No, Dr. Garra has one.

CHAIRMAN GARRA: Just there were several references to this group of where you showed this bar where it showed that cases that the machine got correct and the ones that it missed. Did you detect any patterns in the ones that it missed?

For instance, of the 18 that were missed by the human observers, how did the machine do on those 18? I mean, you showed it for all 80, but I don't know which ones are which.

DR. TOLEDANO: Go ahead.

DR. FREEDMAN: This is Dr. Freedman.

I've looked for patterns and so far have not found patterns in the cases missed or detected by the machine. I'm still looking.

Location does not appear to be a factor in that, and so it may be some other factor that I've not quite understood yet.

In terms of the sensitivity for the lesions, you asked about the actual priors. I don't have that number in my head.

For the smaller lesions, the nine to 15 millimeter, the machine detected 68 percent of them, but I do not remember the number for the priors. I know that it detected them because there was

improvement shown based on the true positives, but I don't remember the number of those.

DR. TOLEDANO: Thank you.

Okay. Well, I'll ask my burning question and that has to do with that 50 percent default. There was a line, and the readers moved a cursor location along the line. They were told to move it to the point that would match their confidence of malignancy, and the FDA has told us that they moved it left or they moved it right, and that the left or right movement correlates very highly with their work-up decision.

And I also just wonder because the FDA notes a bimodal distribution for each of the cancers and non-cancers, what impact that has on the ROC curve analysis.

So I know that there are several issues there. There's a design issue of choosing the 50 percent or why you even have a marker to begin with at all, and then there's the ROC question.

I know there are people who can answer both in the room. So do any other members of the panel have more to say on this 50 percent default issue?

(No response.)

DR. TOLEDANO: No. Okay. Sponsor.

DR. FREEDMAN: This is Dr. Freedman.

First, I want to tell you that we did, indeed, spend a fair amount of time deliberating what historic point we should use, that it was not possible, at least easily feasible within the system that we were designing to record the data to have one that had no starting point.

And so what we considered is starting at the left, starting at the center, starting at the right, starting random. Random also proved to be quite difficult so that in the end we chose 50 percent because we couldn't figure out a logical reason prospectively why that would make any difference.

Now, the second point is they move to the left or the right, and if they moved to the left it meant that it was more benign, and if they moved to the right, they meant it was more malignant, and if you wait a moment, I will find the chart that I think answers that in part.

And that is --

DR. TOLEDANO: If you'll give me the page number, I can find it.

DR. FREEDMAN: Okay. The page number that I would use is 3.B.5.

DR. TOLEDANO: For those who have the full PMA or if it's part of the panel pack, 3.B.5.

DR. FREEDMAN: And what this chart shows is that it has three lines on it, and I'll refer to the one in the middle, which is the mean confidence value, and these are for cancer cases, and you will see that the radiologists as a group used all of the confidence scales on average for making their decision on cancer cases; that there is no abrupt break in the confidence level when one looks at the group as a composite.

In addition, a chart that I did not bring here shows that different radiologists did not have the same cut point. In other words, they did not all cut at 50 percent. They cut between 35 percent and 50 to 60 percent, and so they were not using this as a dichotomy. Different radiologists used different points in deciding whether or not something was likely to be cancer based on their marks of a location.

And so we do not feel they used this as to the left meant benign and to the right meant malignant. We think that each person chose their own operating point, whatever that might be, and if they were to the left of that, they meant it was probably benign, and if they were to the right of that, they

would think it's more likely malignant, but that they did not use the 50 percent chart or cutoff.

I did have a chart also that shows the spectrum of each radiologist, and it indeed does show something similar to this, but with clearly very different thresholds used by different radiologists in that decision.

DR. TOLEDANO: Thank you.

That answers one question, and now for the second question having to do with the bimodality of the distributions and their impact on the binormal model, and he knows that I'm looking at him to answer it.

(Laughter.)

DR. TOLEDANO: Just for those of you in the audience who don't know Dr. Metz, he's probably about the only person in the world who could answer this as well as he can.

(Laughter.)

DR. METZ: Thank you. Thank you, Dr. Toledano.

My name is Charles Metz. I'm Professor of Radiology at the University of Chicago, and I'm here as a consultant to Deus Technologies.

The answer is simple basically. The shape

of the distribution of the responses has no direct effect on the estimate of the ROC curve. In principle, there is some second order effect if the observer crowds his or her responses in such a narrow part of the scale that they start to pass each other, but there's no direct effect whatsoever, and the data, as I read it, wasn't subject to that crowding effect, the second order crowding effect that I described.

DR. TOLEDANO: Thank you, Dr. Metz.

DR. METZ: Thank you.

DR. TOLEDANO: Okay. So we've had some discussion, some general discussion, and what I'd like to do now is turn to the discussion points that were prepared by Mr. Doyle, and Mr. Doyle will have overheads of each discussion point.

So I would now like Mr. Doyle to present to the panel the discussion points that the FDA would like addressed by the panel. Copies of these have been available by the sign-in table outside this room.

So the first one says: please discuss whether or not you believe that the PMA contains sufficient data to conclude that the RapidScreen RS-2000 can reduce observational errors by identifying overlooked cancers on chest radiographs, considering (a) the reproducibility of the computer performance,

(b) the non-location specific versus location specific ROC and sensitivity/specificity results, and (c), in particular, the amount of incremental improvement shown.

Who wants to start the discussion?

CHAIRMAN GARRA: Brian Garra.

I notice that C is the question that we were looking at earlier. The magnitude of the improvement does factor in somehow.

Let me ask the FDA. I guess, Dr. Sacks, you presented the reproducibility results. I personally don't have a lot of experience in determining how reproducibility of a machine when it's aiding a human affects the performance of a human observer. Are there any paradigms or previous experience with this?

And the variability that you were showing was fairly significant. Yet the observer study did show an improvement, it appears. If they had done the machine reading twice, would it have stabilized things, or two or three times?

DR. SACKS: Perhaps the company would also like to answer this.

First of all, let me say that the variability that you saw was not due to the device's

algorithm reading the once digitized film. There was two parts to it. There was a redigitization each time, and then a processing.

Had the same digitization been reprocessed ten times, that would have been almost entirely reproducible almost exactly. So it was the digitization that contains the variability here.

As far as the effect on this, no, we don't have the experience to have an idea of what effect this has on the readings, but it does, of course, add another bit of variability that just makes it harder to dissect things out.

But if we can already dissect things out, that really isn't relevant.

CHAIRMAN GARRA: Okay.

DR. TOLEDANO: So what I hear is that most of the variability is due to the digitization itself.

Before I let sponsor respond, because I see Dr. Freedman ready to hop to that seat, I have a significant issue with the fact that the reproducibility was only evaluated in cancer cases and that there is no evaluation of reproducibility in non-cancer cases.

And I also don't remember off the top of my head if these were only the cancer cases that would

have been similar to the current cancers or of they would have been similar also to the actionable priors.

So more to add or clarify or expand before we --

DR. BERG: Dr. Wendie Berg.

I have one question, and that is related to these issues of reproducibility. Why was the decision made to downgrade the data which was initially digitized at .17 millimeters per pixel, and I think it was ended up at .7 millimeters per pixel? Did you evaluate it without downgrading it to see if it was more reproducible?

DR. TOLEDANO: That's a great question.

CHAIRMAN GARRA: Good point. I saw the .7, and I thought it was a typo.

DR. BERG: Yeah, me, too. I was thinking, "Oh, my God."

CHAIRMAN GARRA: Well, they said .17 in their presentation, and then but I know the manual said .7.

DR. TOLEDANO: Okay, sponsor.

DR. FREEDMAN: This is Matthew Freedman.

I will answer the first part again. What was the exact question in the first part?

DR. TOLEDANO: The exact question in the

first part --

DR. FREEDMAN: Was why we?

DR. TOLEDANO: Why you didn't use non-cancers.

DR. FREEDMAN: Non-cancers. The reason that we used cancers instead of non-cancers in the reproducibility is we know from our internal work that the detections of non-cancerous areas are quite variable, and that the detection of cancer areas are far less variable.

So that if we were to have tested this on non-cancer locations, things like rib crossings, vessels, they tend to be inconsistent in identification. Therefore, we had a problem. What do you really consider to be a gold standard?

And in a clinical setting, the most important thing that we're trying to do is cancer detection reliability, and so that was what we measured for reproducibility.

I might just add that we recognize the digitizer to be a problem, and we use several different digitizers within the company to see if we could find one that did not have these problems, and so far we have used one. We hope that in the future digitizers will become better.

The third point is that increasingly the input will be digital data, and with the digital data input, we should not see anything like this variability. It should be almost 100 percent reproducibility.

And then for the second part of the question as to why the data was downsized, I have some inkling, but I will turn that over back to Xin-Wei Xu to answer.

DR. TOLEDANO: Thank you.

DR. XU: Before I'm going to answer the downgrading, but I want to emphasize what is the variation in our study for repeatability because basically we notice this is due to the digitization variation because digitization gives us presumption process.

Whenever you put a film in, it's actually -- the digitizable was the example of imaging to 2K by 2K with the size 14 by 17, but you never know when the line scan was this time sampling this or next time sampling that. So this is the variation, the cost.

So, however, for already digital image, our algorithm is actually the 100 percent, definitely 100 percent repeatability. So then the algorithm timeless. So in this answer, it's in a digital world.

The image is already being -- digital image is already in there, and should be every time. Whether thousand times or million times you apply the algorithm, the detection always the same.

And now I'm back to answer the question about the downgrading because, you know, in lung cancer detection, which is a nodule, it's a totally different situation for a mammal. In mammal, we deal with micro classification which only minimum error code would be 50 microns, but in cancer we talk about at least a three millimeter or five millimeter or bigger than that.

So it's a totally unnecessary to using that high resolution in terms of computing power or speed. So .7 millimeter is sufficient enough. So that's why, the only reason we downgrade from original 2K by 2K or the one we apply our algorithm, either downgrade to a pixel size of .7.

DR. BERG: As a follow-up question, are you sure that that's not affecting future analysis by doing that?

DR. XU: We basically don't see that happen. Yeah, basically the feature we deal with, either most logical feature like shape, size. These are attracted from the reasonable large site. At

least in our case, more than nine millimeters. So .7 millimeters are really -- it's a not big concern for us.

DR. TOLEDANO: Thank you.

Do we have further discussion of this point before we make a decision about this discussion point?

We actually need to make a decision about this discussion point before we can move to the next discussion point. So what do people think? Do we need to discuss more or are we ready for a decision on the discussion point?

Dr. Mehta.

DR. MEHTA: Can I ask a clarification?

Minesh Mehta here.

If I understand correctly then, the reproducibility issue, which is the point we're talking about right now, we are at this point of understanding, that the vast majority of the error in producibility comes from the process of digitization, and that if the chest X-ray were entered into the system as a digital radiograph to begin with, this would go away.

Have you done that experiment? And is the system capable of inputting data directly digitally?

And so what are the results?

DR. XU: Yes, the answer.

DR. TOLEDANO: Please can you say whether or not you believe that the PMA contains sufficient data to conclude that the RapidScreen RS-2000 can reduce observational errors by identifying overlooked cancers on chest radiographs considering the three items listed below?

And we'll start with the new guy, Dr. Smith.

DR. SMITH: Oh, boy.

(Laughter.)

DR. TOLEDANO: I know. I'm not supposed to start with you. Here you go, Dr. Garra.

CHAIRMAN GARRA: That's what I'm known as in the department. "He's the oldest member in this room."

I would say that I think that the answer is whether or not -- I think that the PMA does contain sufficient data to conclude that the RS-2000 can reduce observational errors.

I think there's one factor that probably wasn't listed there that probably should be, and that's training because we did see some reductions in performance by some of the observers, and the evidence

is tending to point towards training issues as far as why they reduced their performance. They were apparently using it inappropriate or somewhat inappropriately.

So that will have to be something that careful attention to is paid -- careful attention is paid to.

DR. TOLEDANO: Thank you, Dr. Garra.

Training is one of the further discussion points as well.

Okay. Dr. Smith, are you ready now?

DR. SMITH: I think so. I would have to agree. I think it can reduce observational error, taken just as a yes or no question.

DR. TOLEDANO: Thank you.

Dr. Mehta.

DR. MEHTA: Alicia, let me just be sure I understand the question correctly. Are we talking about subpoint A or the entire question?

DR. TOLEDANO: We are talking about the entire question, whether or not you believe that the PMA contains sufficient data to conclude that the RS-2000 can reduce observational errors by identifying overlooked cancers.

And when you answer that question or when

you're forming the answer to that question, you should consider at a minimum the three subpoints listed below.

DR. MEHTA: I would have to say no, specifically because I'm not convinced that the incremental improvement is of sufficient value.

DR. TOLEDANO: Thank you, Dr. Mehta.

Dr. Harms.

DR. HARMS: I would agree.

DR. TOLEDANO: Thank you.

DR. BERG: I would like to see more discussion of Points B and C.

DR. TOLEDANO: Okay. Should we discuss Points B and C more before we come up with a final answer to the question?

I guess we will. Dr. Berg.

DR. BERG: Well, I think, in particular Point B is curious to me because I think one of the issues really is this location specific analysis, and actually the way the question is worded, it says location specific ROC, which I thought we couldn't do, but the question in my own mind is if you randomly scattered marks on a film and submitted to radiologists, they're going to look at that film a second time in and of itself how much of the benefit

is just from that.

In the back of my mind that's what I'm trying to get at, and I'm not sure that I have a really good answer to that. From the statistical presentations of the FDA this morning, it would appear that at least half of the benefit was location nonspecific. In other words, it probably was just that issue of looking at the film a second time.

So I would like that to be addressed.

DR. TOLEDANO: Do other members of the panel have insight into that concern? Perspectives on that concern?

Sponsor.

CHAIRMAN GARRA: Well, Actually before you make the comment, Brian Garra here.

Definitely location specific. I think you could have done location specific on ROC, but -- I know -- but even when you did the location specific analysis, it reduced performance, but it didn't eliminate it

DR. TOLEDANO: Okay. I'm going to let sponsor respond, and then actually we're sort of running shy on time. So we've got four more to discuss in the next half an hour. So I will request that sponsor keep comments and replies brief.

Sponsor.

DR. FREEDMAN: We believe that the ROC area is the best data for this decision. We did not design this to test people's ability to detect the correct location as a measure of the performance of the machine, and the reason is that we knew that these cases had more than one area of positive signal, of potential lesions on the film.

Therefore, if one were to say, "Do a study to determine whether or not someone can detect location specific information," we would have allowed them two or three choices of location.

That was not the purpose. The purpose was to learn exactly the question that you asked the question about, which is: what is the effect of a machine negative on the performance?

Now, a machine negative means that the machine has marked a location on the film, but it is not the cancer, and what happened in those cases you can see by the very slight decrease in specificity, which means that, indeed, occasionally when there was a mark in the wrong location that the radiologist did respond to that mark, but that that was relatively infrequent.

The amount of gain that we saw I don't

think can be explained by the presence of those random marks. Also, the radiologists were not responding to nothing. They would have been responding to something that was identified on the film.

Many of them can be eliminated, rib crossings and so on, but some of them are scars. So fundamental problems.

One, we didn't ask them to do the task that the FDA is saying we did. We were using that data for something different and, therefore, the results are predictably inferior.

The second thing is that we know that the false marks, if there's a real lesion there -- someone will say, "That is not a scar. That's a cancer." And as I said, in the Hopkins study, they called back 25 people for every cancer seen. So it's not surprising that the radiologists with the aid of this would pick up very subtle lesions that were not cancer.

Does that answer your question?

DR. BERG: Sure. Thank you.

DR. TOLEDANO: Thank you.

Actually that's about all of the time that we have to discuss Point B and C because we really do need to move, and I do apologize if I did not pace the earlier discussions quickly enough.

The second discussion point -- is there somebody who's supposed to press a button? -- if you conclude -- ta-da.

So we did not come up with a conclusion, a clear conclusion either way. We were a split panel. So we will just end up having to discuss whether or not we believe that the PMA contains sufficient data to conclude that this can be done without unacceptably increasing the number of patient work-ups.

So I guess the idea is given that we can reduce the observational errors by identifying overlooked cancers, can we do this without unacceptably increasing the number of patient work-ups?

MR. SEGERSON: Dr. Toledano.

DR. TOLEDANO: Yes, Dr. Segerson.

MR. SEGERSON: Let me clarify something. These issues are really meant to walk you through all of the issues that we identified in our review. We're not really looking for a conclusion on each point right now. We're going to be asking you to vote on the approvability of the PMA later.

DR. TOLEDANO: Right.

MR. SEGERSON: But this is just an exercise to get the discussion on the floor, and

whether or not you actually draw a panel conclusion on each point is not necessary.

DR. TOLEDANO: Thank you, Dr. Segerson.

Actually the reason I had sought it on the first point is that the second one depended on whether the -- depended on the conclusion to the first, but, yes, I do remember that we just basically discuss and move on and discuss and move on.

So that's what happens for anybody who hasn't been to these things before. We discuss and move on and discuss and move on, and thank you, Dr. Segerson, for reminding me of that.

So would we like to discuss or would we like to move on?

(Laughter.)

CHAIRMAN GARRA: Dr. Garra here.

I think we have discussed this. I don't know the answer to that question. So I guess -- because I don't know what unacceptable is and I don't know that there's sufficient data. I don't know how the other panel members feel on this one. This is a tough question.

DR. BERG: Dr. Wendie Berg.

Yeah, I have trouble answering this question also because I think there's a relatively

small number of normals, and you know, are we going to be proposing to use this in an in-patient setting where there are going to be a lot of other issues going on in those patients' chest X-rays or not?

All of these questions are hard to answer, I think, from the data we have.

DR. HARMS: Steve Harms.

This is not a prospective trial. This was a selective group of films to test the machine, and this is more of a clinical question. You know, if we start using this in clinical practice, are we going to have an unacceptable number of false positives generating work-up?

I think probably the way I view this is that what is unacceptable is doing a chest CT an unacceptable outcome. I don't think it necessarily is. In fact, we're thinking about doing screening CTs anyway. So the down side risk of this is nonacceptable, and I would be willing to tolerate a fairly high number of false positives.

DR. TOLEDANO: Thank you, Dr. Harms.

Further contribution to this discussion? Dr. Mehta.

DR. MEHTA: Minesh Mehta here.

I just have a scenario that, again, I

think it's a hard question to answer. I don't even think we can come up with it, but here's a scenario.

Let's go back to the average hospital. You take the next 10,000 chest X-rays that occur in the average hospital. We detect five lung cancer cases with the help of our radiology team. We know from the statistical data that was presented to us at the lower limit of statistical improvement with CAD is 1.96 percent or, say, two percent, which means that with this we'll detect 5.1 cases for the 10,000 chest X-rays.

In other words, for every additional case that we'll detect, we'll process 100,000 films. That's just processing the films. I don't know what's happening to the patients.

So unless we have those kinds of numbers -- those are just numbers I made up as we went along -- but unless we have numbers like that, we can't answer this question.

DR. TOLEDANO: Dr. Smith.

DR. SMITH: And I guess just to echo, I think it is going to increase the number of patient work-ups. It just really hinges on what you consider unacceptable, and that almost gets into areas of cost-benefit that we're not considering.

MR. SEGERSON: Dr. Toledano.

DR. TOLEDANO: Yes.

MR. SEGERSON: I think the cost or risk that we're talking about, the unacceptability has only to do with patient risk. This came out once before when we were talking about risk-benefit. It's really risk when we say cost.

But if it's really risk-benefit, then we're talking about the patient.

DR. TOLEDANO: Okay. So thank you, Dr. Segerson, focusing in on the risk to the patient.

Does anybody have more to contribute on it? I know people have said things about the particular patient and the risk to the patient already.

CHAIRMAN GARRA: Is that risk to their pocketbook or risk to their body?

(Laughter.)

MR. SEGERSON: Their health.

CHAIRMAN GARRA: Because the number then wouldn't count. Even one would be not too good, right?

DR. TOLEDANO: Dr. Mehta.

DR. MEHTA: Minesh Mehta here.

I do have something to say about risk to

health. Even if a CT -- let's say we all assume for a moment that the low dose radiation that arises from CT is completely harmless in these patients, if the CT is causing no harm. The psychological harm of having to be worked up for a cancer in a large population of patients where a very tiny fraction of them will be found to have cancer might be substantial. And that's a health risk, the psychological harm.

DR. SMITH: Also, if I may -- John Smith -- if you're giving contrast to, say, your 10,000 patients, there is a defined risk with that, even low osmolar contrast materials. About 30 out of 100,000 will have a serious reaction.

DR. TOLEDANO: Dr. Harms.

DR. HARMS: Typically the screening CTs won't be done with contrast.

DR. BERG: Right.

CHAIRMAN GARRA: Because there will just be a nodule detection run.

DR. BERG: Right.

CHAIRMAN GARRA: Yeah.

DR. TOLEDANO: Thank you.

I was seeing Dr. Metz raising his hand. So let's give him 60 seconds.

DR. METZ: Thank you.

Charles Metz.

I'd like to make an ill prepared comment because I can't find the numbers in my notes, but with regard to the increase in finding cancers, if I recall correctly and perhaps someone can point to page, the gain in the area under the ROC curve was on the order of two percent, and that the gain in sensitivity was on the order of seven or eight percent. Can someone point me to that?

DR. TOLEDANO: I believe it was six percent.

DR. METZ: Okay. Six or seven percent. So it's a lot bigger.

Whether that's acceptable, of course, is another question, but it's a lot bigger than the two percent.

DR. TOLEDANO: Thank you.

MR. SEGERSON: Can I ask Dr. Kondratovich to come to the microphone? Do you mind?

DR. TOLEDANO: I don't mind. Go ahead.

DR. KONDRATOVICH: Marina Kondratovich, biomedical statistician.

If you remember that 6.5 is the point estimate of sensitivity at confidence interval was relatively big. Therefore, area under the curve, of

course, more reliable characteristic in this situation. That trait of if you would like to use area under the curve, then you have more sensible tool. If you would like to use point estimate like sensitivity/specificity, but the confidence interval very huge, and even you can see that for comparison with independent without computers, the confidence interval contains zero.

DR. TOLEDANO: Yes, thank you.

Okay. Further discussion of this second discussion point? And remember, we're just focusing in on whether it is an unacceptable risk to the patient's health.

(No response.)

DR. TOLEDANO: Okay. I see no further discussion of this point from the panel.

Discussion point number three. Please discuss whether the labeling of this device -- oh, this is always the one that takes a really long time -- please discuss whether the labeling of this device, including the indications for use, is appropriate based on the data provided in the PMA. Consider as a minimum:

(a) The ability to detect solitary pulmonary nodules; Oh, you all have it numbered.

Two, the ability to detect more, what size range, solitary pulmonary nodules using the device than when not using it;

Three, the ability to reduce the likelihood of missing small lung cancers, most of which are early stage cancers; and

Four, the target population, for example, age, smokers versus non-smokers, pack-years, et cetera.

I would say discuss amongst ourselves, but we're here in public for a reason. So would anybody like to start the discussion? I will cold call. Would anybody like to start the discussion?

Dr. Mehta.

DR. MEHTA: Minesh Mehta here.

I'll start in reverse order. I think the target population might be a straightforward one to handle because we know what target population is part of the study. So let's look at what target population was not part of the study, which means the PMA does not contain data on that target population.

It does not contain data on predominantly non-smokers. It does not contain data on females, and it does not contain any pack-year related data. So I think it's obvious those three elements are missing.

They're all adults in this study. So there's no data on none of those.

So those are four patient groups for which we have no data in the PMA for which, therefore, we cannot make conclusions.

DR. TOLEDANO: Thank you, Dr. Mehta.

CHAIRMAN GARRA: I have a question for the FDA members. We did have data presented on women, not part of the clinical trial. Does that factor in? Is that data that we officially have that we can use in our determination?

DR. TOLEDANO: Yes/no answer from the FDA?

I believe, Dr. Garra, you're referring to the slide that was shown?

CHAIRMAN GARRA: Un-huh.

DR. TOLEDANO: Yes, an overhead that was shown by Dr. Freedman.

DR. SACKS: That was not part of the PMA.

DR. TOLEDANO: Okay. Thank you, Dr. Sacks.

Further questions?

I also wanted to note as far as I could tell from the PMA, these were all men age over 45 years in the sample? Simple yes/no answer.

DR. FREEDMAN: Yes.

DR. TOLEDANO: Yes. Thank you, Dr. Freedman.

So in terms of the target population, we have men. We have data from a clinical ROC study for films collected 25 years ago for men over the age of 45 who were heavy smokers, and I think that summarizes the point that Dr. Mehta was making.

Further discussion of any of the other points, the other subpoints?

(No response.)

DR. TOLEDANO: I see no further discussion of the other subpoints. Okay. Dr. Segerson.

Okay. I have two other discussion points. Would you like to put the indications up now or --

MR. SEGERSON: Well, this focus on labeling, including the indications for use, I didn't know if it might help to actually display the indications again, unless you wanted to just look at your own copies.

CHAIRMAN GARRA: Yeah, I think the issue here is that labeling is fairly extensive.

MR. SEGERSON: Yes.

CHAIRMAN GARRA: And to determine where these have to be applied takes more than a couple of seconds.

MR. SEGERSON: If you make some comments on the indications for use, then we can subsequently probably extrapolate from that to the rest of the labeling after the panel meeting.

DR. TOLEDANO: Okay. That's helpful. Would we like to see the indications or would we like to refer to our own copies?

We'll just refer to our own copies. Okay. So computer aided -- it's a CAD system to identify ROIs on digitized frontal chest radiographs. So that's the first thing, frontal chest radiographs that may have features associated with solitary pulmonary nodules. So that's your second item, solitary pulmonary nodules from nine to 30 millimeters in size. So there's your third item, nine to 30 millimeters in size, which could represent early stage lung cancer.

The device is intended for use as an aid only after the physician has informed an initial interpretation of the radiograph. Thus, the device assists the physician in identifying areas containing a potential lesion that previously may have been missed.

And I know we've discussed the second two sentences earlier in this discussion period. So let's focus on the frontal, the SPNs and the sizes. Do

people have comments, questions about those?

Dr. Mehta.

DR. MEHTA: One minor comment on the frontal. Minesh Mehta here again.

Of the 250 X-rays, how many were posterior anterior views and how many were anterior posterior views?

DR. TOLEDANO: Okay. So short answer from sponsor? Number or a symbol. I don't know if that's the --

DR. FREEDMAN: I don't know.

DR. TOLEDANO: Thank you.

DR. FREEDMAN: But I can say that -- this is Freedman -- they were primarily PA, but they're not labeled, and if they're done at six feet, I can't tell the difference.

DR. TOLEDANO: Thank you, Dr. Freedman.

Dr. Mehta.

DR. MEHTA: Minesh Mehta here again.

If they were primarily PA, that's what the indications should state, PA chest X-ray.

CHAIRMAN GARRA: I would disagree with that. If they're partly PA and partly AP, then I think either is acceptable.

DR. TOLEDANO: Thank you, Dr. Garra.

Dr. Berg.

DR. BERG: Wendie Berg.

One question, and that is that as I recall, but I can't lay my hands on it right this second, there was no net benefit with the device with cancers over a certain size. Do we need to address that in the labeling?

DR. TOLEDANO: Go ahead, Dr. Garra.

CHAIRMAN GARRA: My comment, again, would be that, yeah, they didn't show a benefit and it was usable in those devices.

DR. BERG: Right.

CHAIRMAN GARRA: So you could use it on them and it didn't hurt, but I think there has to be a little extra here or somewhere in the labeling that is more clear about where the benefit really lies with this instrument.

DR. TOLEDANO: So I'll actually make a little comment on that, which is that the primary hypothesis was for the nine to 30 millimeters, and for that overall size range, there was a benefit shown.

Now, what happens when you go into the subgroups is that you could see a benefit because you see a large benefit, or you could see a benefit because you have a lot of lesions in the subgroup.

You could not see a benefit because a benefit doesn't exist or you could not see a benefit because you don't have a sufficient number of lesions in that subgroup.

And because these are subgroup analyses, I think it's particularly dangerous to say that we didn't see it. When we see it in the overall, when we say nine to 30 and we look overall and we see it, I think that's fine. I think as we get into the subgroups, if we're making claims, yes, the claim for nine to 14 or nine to 15 is valid. The claim for 15 to 19, that was not statistically significant, and we would need to partition out whether that's because of a sample size or because it's just not significant.

And then for the 20 to 27.5, that was just a wash. So -- and I see Dr. Sacks nodding. So I hope that wasn't too much of an odious or presumptive comment for me to make.

Dr. Garra.

CHAIRMAN GARRA: I was just rather -- since the numbers are going to be a little uncertain, I think most of the lesions were in the smaller size ranges.

DR. TOLEDANO: Yes.

CHAIRMAN GARRA: Is that right? I think so.

DR. SACKS: Thirty-five were in the smallest, 25 were in the middle, and 17 in the highest.

CHAIRMAN GARRA: Okay, yeah. So of the three classes, the one where they got nice results were in the small size range where they have fairly large numbers.

So if we were going to add material to this, I would say that it's primarily useful as an aid in the smaller nodules, nine to 15 millimeters and leave it at that because we don't know whether the data in the larger nodules are not significant because of the smaller numbers or smaller change or a combination of the two.

DR. TOLEDANO: Thank you.

Further comment on this point? This is the size.

Okay. So we've discussed frontal. We've discussed size. Did we discuss the fact that they're SPNs? Did we want to discuss the fact that these are SPNs, solitary pulmonary nodules?

Everybody is tired and wants to go home. We still have two discussion points left.

Okay. So I don't see any further discussion of the indications. Let's move to our next

discussion point. This is the fourth discussion point that has to do with the training.

Oh, did you skip one? Is there another one or did it get deleted?

DR. PHILLIPS: No, it's there.

DR. TOLEDANO: It's there? Okay. I'll start reading it while it's located.

Based on the information shown in the PMA, were the film readers sufficiently trained in the use of the device before the with CAD readings were made? And what implications does this have for training or prospective users of the device if it is approved.

Now, I know there's bound to be comments on this because there were eight films, three of which contain cancers. But not being a radiologist myself, can I have somebody else begin the comments?

DR. BERG: I'm Dr. Wendie Berg.

I guess I'm concerned having done some reader studies. That should really suffice in training people, and they should really have enough common sense to not change a true positive to a false negative.

So I'm a little concerned by the fact that there were some that did. I don't think the training is going to be changed that much in practice. Any

radiologist is not going to be spending a whole lot of time wanting to be trained on this device.

With experience, it might improve, but I think there is potential down side that probably needs to be reflected in the labeling.

DR. TOLEDANO: Thank you.

Further comments?

I'm just reading that I misread my agenda, and we were supposed to stop five minutes ago, and I thought we had another ten.

Are there any further comments on the training? Okay.

MS. PETERS: I have a question

DR. TOLEDANO: Yes.

MS. PETERS: This is Marilyn Peters.

I was just wondering on the training did you find out why people changed their mind. Is it because they were just being independent or they didn't understand the instructions or what?

DR. FREEDMAN: This is Dr. Freedman.

We had a post session interview and a form, and the people in general, in fact, uniformly really liked the system and said they thought it was very helpful, and it's only when we did the data analysis that we found these problems.

DR. TOLEDANO: I guess I would have another question, which has to do with when we tend to do reader studies, ROC reader studies, we tend to pick on the experts, and that's just because those of us who are putting together the studies tend to know the experts.

This particular study was specifically designed to use people in the field, which has a wonderful implication for the clinical use, but I wonder what role did that play in this switchability and susceptibility to switching.

Anybody want to make any -- do other panel members have any insight into that?

CHAIRMAN GARRA: That was one of several issues that I was wondering about. Nancy, I think you said that the -- you chose 240 cases because that's the number that a person could comfortably do in a half a day. With filling out all of the forms, is that correct? Did I misunderstand you there?

The only reason I bring this up at this point is because somebody is more likely to maybe use the machine instead of reading the film if they're rushed.

DR. TOLEDANO: Go ahead, Dr. Freedman.

DR. FREEDMAN: Oh, I misspoke before. It

was two half days with a split in the middle of the days. So it was one day as long as they took to complete the task within that period of a day.

DR. TOLEDANO: Thank you.

Go ahead.

CHAIRMAN GARRA: I'd just like to comment I still think that's a significant burden for a reader, and I could see where they might be rushed if they have to -- I don't know how easy it was to do the -- like if they clicked the wrong button and they had to fill out the little form.

Yeah, maybe we could get a comment on that.

DR. TOLEDANO: Comment, no comment, or a comment that is not difficult at all?

We need a name and a microphone actually.

DR. KHAZAN: Ron Khazan.

It was a slight burden, but it was doable.

If I may answer one question I heard a comment on about the two minutes per nodule to assess, it's nothing near that. You look at five ridiculous flags from the computer, and it may take ten seconds to dismiss them all.

If there's a reasonable one, you might spend, you know, half a minute looking at it.

DR. TOLEDANO: Thank you.

CHAIRMAN GARRA: Just another comment. I did note, however, that several times the machine flagged regions that were cancer, and they were dismissed.

DR. KHAZAN: Right, right.

CHAIRMAN GARRA: And that's another training issue maybe where with more extensive training on trying to detect patterns.

DR. TOLEDANO: Okay. I'll recognize Dr. Freedman.

DR. FREEDMAN: Thank you.

This is Dr. Freedman.

I'd like to propose this as a challenge. How many of you have ever detected a lung cancer on a chest X-ray 15 millimeters or less and how consistently? These are cancers that are at the threshold of what most people consider detectable. They are smaller in size than what has been found as the average size of previous screen trials. This is the size of cancers that were previously missed. They are very hard to see, some of them.

And, therefore, it is not surprising that even with it circled, someone could look at it and say, "I don't think there's anything there."

But the expert panel confirmed that there was something there because we knew the location where it was one year later.

So the fact that the radiologist did not always accept the computer is not surprising to me.

CHAIRMAN GARRA: Again, a comment. With additional training, they might reset their decision threshold to accept more of those, of course, perhaps at the cost of more false positives.

DR. FREEDMAN: This is Dr. Freedman.

I agree completely, and I actually am working on a training CD.

DR. TOLEDANO: Thank you, Dr. Freedman.

Dr. Wagner has a --

DR. WAGNER: A quick comment.

DR. TOLEDANO: -- a quick comment. Okay.

DR. WAGNER: Bob Wagner from the FDA.

No one touched my comment whereby I pointed out that we were just talking about the smaller lesions, that over half of the readers scored in the high 80s or the low 90s for the smaller lesions with the use of CAD.

It's a remarkably high score card I would consider for this. We think we see small effects here, but this is rather large.

DR. TOLEDANO: Thank you, Dr. Wagner, for that very insightful comment and useful comment.

Further discussion of this fourth discussion point before we take four minutes to discuss the fifth discussion point?

Okay. Does somebody want to push that button for me?

The final discussion point: do the above or any other issues not fully addressed in the PMA need resolution before the PMA is approved, require post market surveillance or suggest a post market study?

This is where we always end up with this panel, isn't it? Any ideas about post market surveillance or post market study, requirements for, recommendations for, necessity for?

CHAIRMAN GARRA: I would only comment that if the FDA approves it to include groups other than the groups that Dr. Mehta was talking about, for instance, if it's approved and there's no mention made that it was not tested on women, that you might have to have a post market study to confirm that it is, in fact, helpful there, although the evidence that we didn't see from Matthew suggests that it probably will be.

DR. TOLEDANO: Thank you, Dr. Garra.

Dr. Harms.

DR. HARMS: Just a comment on the need for a post market study. Usually that's done when there's a question about risk, and this has almost no risk, and therefore, I think the need for a post market study is nil.

DR. TOLEDANO: Thank you, Dr. Harms.

Do other members of the panel have further comments or agreeing or opposing viewpoints about the need for a post market study?

CHAIRMAN GARRA: I would just comments as far as the post market study though. I think that, I mean, we're supposed to use both risk and benefit, and although it's more obvious that you need to do a post market study if there's risks involved, I think that if you allow a device to be used and market it for a group for which there was no data, then a post market study may be indicated.

DR. TOLEDANO: Dr. Garra, do you think that, for instance, on the question of women or non-smokers or these populations, do you think there's a need for a post market study or do you think that post market surveillance might be sufficient?

CHAIRMAN GARRA: Surveillance would be

fine. The issue here is what occurs in the labeling, what occurs in the indications for use, and the FDA will keep a tight grip on that, I'm sure.

Of course, it will be used in these other groups. It will be used off label. So we're not trying to restrict its use. It's just a matter of what they advertise it for.

DR. TOLEDANO: We have one minute for further questions, and then I'm supposed to let everybody take a break. Does anybody want to ask any further questions?

Oh, we have a question from the sponsor. Okay. Come on.

DR. FREEDMAN: I think it's important to understand one problem with anything other than surveillance, and that is that there is no lung cancer screening clinical type protocol in the United States now using chest X-rays. There is CT randomized to chest X-ray, the POCO study. We have approached them to see whether we could get our device incorporated into that as a clinical arm or as independent of the clinical arm, and they have said no, because the study design is set.

We have approached the lung screening study run by NCI. Those are formulated and set, and

it is very difficult for a sponsor to then set up a clinical trial.

It is very different than in mammography where breast screening is done routinely. Lung screening with a chest X-ray is not done routinely. CT screening is still experimental, and at least the formal programs are still under statistical control as to what can be used and added to them.

Surveillance is fine. A post market study may be very difficult and expensive for a small business to handle.

DR. TOLEDANO: Thank you, Dr. Freedman.

Dr. Mehta, I'm going to let you go real quick.

DR. MEHTA: Minesh Mehta here.

On the basis of that comment from the sponsor, I actually do have a question now. I certainly appreciate and understand the difficulty of mounting a screening study, but since the labeling of this device does not have the word "screening" anywhere, can we focus on what the labeling is all about? That's as an adjunct.

If a radiologist can go through 200 chest X-rays, if you wanted to look at 10,000 sequential, unselected chest X-rays, it would take you 50 days.

Could you do a post market surveillance in 50 days, look at 10,000 X-rays and answer my original question, how many more cancer cases did you pick up, because then we'd have a cost-benefit answer to this right away?

DR. TOLEDANO: Sponsor, I'll let you do a quick yes/no.

(No response.)

DR. TOLEDANO: Okay. I'll let you do an "I don't know." He's just turning red.

I think we've had actually a very active discussion. I think almost everybody has participated. Pretty much everybody has participated very vocally.

So I would just like to say thank you to everybody for all of your comments and for all of your contributions to this discussion, and now I'm supposed to turn this back to Dr. Garra. I just have to find the part in my script where I turn it back.

CHAIRMAN GARRA: That's sufficient.

(Laughter.)

CHAIRMAN GARRA: I'm taking control again here.

Okay. The question I have for the panel right now is we normally would have a 15-minute break